



Virtual in-ear microphone for in-vehicle noise control based on array technology and modified zero point attraction LMS algorithms

Mirjana ADNADJEVIC¹; Dick BOTTELDOOREN²

Ghent University, Belgium

ABSTRACT

Measuring in the ear of an operator of a vehicle or heavy duty machine may be useful in several applications such as noise assessment, adaptive communication, headphone-less binaural sound reproduction, or active noise control. However, wearing in-ear microphones is not very comfortable for the operator. Therefore, we investigated the possibility to replace these microphones by virtual microphones based on an array of microphones optimally located in an enclosure. Several modifications to LMS based algorithms for sparse MISO system identification were proposed and adopted at the one hand to eliminate as many of the filter taps as possible to reduce the computational load, and at the other hand eliminate as many microphones as possible to reduce hardware costs. Identification methods are evaluated through the simulations based on the experimental measurements performed in the real size driver's cabin constructed for the purpose of active noise control (ANC). Forty six microphones were initially positioned at the cabin roof, while virtual location is chosen to be ear of the seated observer. It is shown that these algorithms indeed reduce the set of microphones, still providing acceptable fitting performance when compared to the reference NLMS technique where all available microphones are used. Sensitivity of converged coefficients to natural head movements is evaluated for the head rotations of 20, 45 and 60 degrees in horizontal plane.

Keywords: LMS algorithm, sparsity, microphone array, ANC, Active Noise Control
I-INCE Classification of Subjects Number(s): 38.2, 74.7, 14.4.7

1. INTRODUCTION

The quality of the working environment and in particular the acoustic comfort has been a topic of concern for many years and is increasingly important in industrializing areas around the world. Indeed, excessive noise exposure may lead to occupational noise-induced hearing damage and/or lower job efficiency (1). With increasing automation and more complex machines being introduced, the human operator is now often located in a more pleasurable environment shielded from the noisy part of the machine and noise control moves from simply reducing noise levels to sound quality engineering. Numerous active noise control algorithms have been proposed in literature for this purpose. They are mostly tuned for local control corresponding to the small area around the operator, as global control is often not efficient due to the complex sound field in the enclosed spaces resulting in high hardware costs (2). However, the main constraint of traditional local control approaches is the limited zone of control located around the error sensor. For best performance, the error microphone should be placed in the ear of the operator, assuring the controllability of the actual exposure. Since wearing in-ear microphones or microphones very close to the ear is usually impractical, virtual microphone techniques are proposed that can shift the zone of control to the desired location which is remote from the physical sensor (3). The use of virtual in ear microphones is not limited to active noise control. They could also be used for tuning headphone-less binaural sound reproduction with applications in music and communication. In particular making communication signals appear virtually from a given location may be useful for informing operators of machines or vehicles about their surroundings.

Assessing the acoustic field at a specific location remotely requires preliminary identification of the transfer functions between the microphones that will be used and the physical microphones placed at the virtual locations. After identification is performed, these are removed from the virtual locations and the transfer functions are applied to the remaining microphones to obtain the signal at the virtual location. Different virtual

¹madnadge@intec.ugent.be

²dick.botteldooren@intec.ugent.be

sensing algorithms have evolved depending on the identification method applied. Some of them are based on FIR and IIR black-box modeling, while others utilize state-space modeling methods and different derivatives of adaptive LMS techniques (3).

In this study we are particularly interested in in-ear virtual sensing techniques for Active Noise Control (ANC) in the driver's cabin of heavy duty machines. The possibilities for using a microphone array positioned at the roof of the cabin for sound field estimation at the ear of the driver is investigated. The normalized LMS algorithm is used as a reference approach to derive weights for a Multiple Input Single Output (MISO) system due to its simplicity, robustness and low computational cost (4). This reference algorithm performs well with a large array of microphones and a sufficient number of filter taps. However, the hardware cost could be considerable. In addition to the microphones, also the signal processing power needs to be increased to allow performing the microphone virtualization in real time. Moreover, there is a risk of over-fitting to the virtual microphone positions and source locations used during the identification phase.

To tackle the first problem, several modifications to the NLMS algorithm are proposed that apply l_1 relaxation, common in compressive sensing, promoting the sparsity in the taps during the identification procedure:

1. WZA-NLMS (Windowing Zero-Attracting NLMS),
2. PWZA-NLMS (Power Windowing Zero-Attracting NLMS),
3. Modified RZA-NLMS (Re-weighted Zero-Attracting NLMS),
4. PRZA-NLMS (Power Re-weighted Zero-Attracting NLMS), and
5. RW-RZA-NLMS (Roulette Wheel Re-weighted Zero-Attracting NLMS).

The first two algorithms both increase the number of near-zero coefficients of the system by applying a constant zero attractor to the coefficients when they are in the pre-defined amplitude range [a-b]. The second algorithm in particular applies the WZA-NLMS selectively only to the microphones whose impulse response does not satisfy certain power constraints, and leads to the minimal set of microphones still fulfilling fitting requirements. The third algorithm, a modification to RZA-NLMS, employs a re-weighted zero attractor for different microphones instead of filter taps, relaxing the l_1 norm penalty for those with more significant impulse responses according to the pre-set absolute value criteria. The advantage is in on-line selection of optimal microphones with increased number of eliminated microphones, while preserving the quality of the identification. To further improve performance, two extensions of RZA-NLMS are considered: 1. PRZA-NLMS following the same idea as PWZA-NLMS and 2. RW-RZA-NLMS which uses roulette wheel rule to make a random choice of the microphones which are not to be affected by zero attraction. Advantage of the later approach is in the ability to eliminate two similar microphones. The choice of the identification parameters (amplitude attracting range, absolute value and power criteria) is investigated in the paper.

The second problem is addressed by investigating the sweet spot based on measurements in a mock up at real scale of a driver's cabin.

2. ADAPTIVE LMS-BASED IN-EAR VIRTUAL MICROPHONE

2.1 Conceptual scheme in the context of ANC

An LMS based virtual microphone uses any modification of the standard LMS algorithm to obtain an estimate of the signal at a fixed virtual location by combining signals from the remotely located physical sensors. For integration in an ANC application, the virtual location is the ear of the seated observer and signals at both physical and virtual locations are referred to as "error signals" representing the interference pattern of the primary and secondary (controlling) field. According to (3), for more complex sound fields, as is the case in a cabin of an agricultural machine, one unique set of weights which is optimal for both sound fields (primary and secondary) is hard to obtain. This is especially true in the near field of secondary sources where these fields differ significantly (3). It is suggested in (5) that optimal weights for the estimation of primary and secondary fields should be found separately. Once identified, weights are then integrated with the algorithm for ANC following the block scheme shown on Figure 1. During the ANC, primary and secondary fields are separated out at each physical microphone using the knowledge on the secondary transfer functions (\tilde{G}_{pu}) identified a priori. Identified weights (w_p and w_u) are then applied to the corresponding fields for the reconstruction of the error signal at the virtual location (\tilde{e}_v). Due to the lack of time, only weights for the secondary field are derived by applying different NLMS based identification methods. Each of the method is explained in following subsection, while the results are presented in dedicated section. For the simplicity, weights related to the secondary field (w_u) are further in the text substituted with w .

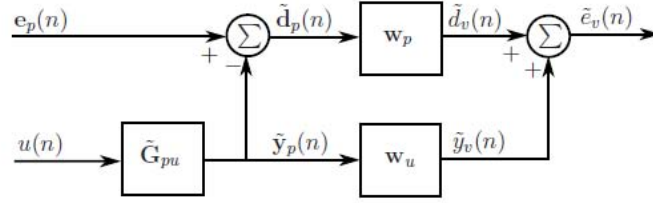


Figure 1 – Block diagram of the adaptive LMS based virtual microphone technique, adapted from (3).

2.2 Identification methods

Several identification methods featuring basic LMS properties are proposed to calculate physical microphone weights required for virtual sensing. Although evaluated in the driver's cabin and for the secondary field case, their basic principles could be easily extended to any kind of acoustic field and environment. Each identification method is adjusted for MISO system identification with $N_p = 46$ number of inputs (remote physical microphones) and one output (virtual location at left/right ear). Moreover, adaptation of the weights is adjusted towards the optimal solution for all $K = 3$ secondary field sources. During the identification stage, sources are independently excited with band-limited white noise, while unknown coefficients are estimated using the N_p signals from the remote microphones and desired signal from the microphone temporarily placed at virtual location.

NLMS algorithm is used as reference case as it minimizes the square difference between desired signal at virtual location and weighted summation of the signals from the array microphones. Its cost function and weights update rule are given in as:

$$J = \sum_{k=1}^K e_k^2 = \sum_{k=1}^K \left\{ y_k - \sum_{j=1}^{N_p} \sum_{i=1}^M x_k(i, j) \cdot w(i, j) \right\}^2 \quad (1)$$

$$w_{n+1}(i, j) = w_n(i, j) + \sum_{k=1}^K \frac{\mu}{\sum_{i=1}^M x_k^2(i, j)} \cdot x_k(i, j) \cdot e_k$$

where M (50) is a number of filter taps assigned to the each microphone in the array, $x_k(i, j)$ signal-tap at the j -th remote microphone when excitation was k -th source. Parameter μ is set to 0.0000001, while different step-size is applied to each array microphone, adjusted according to its signal power, ensuring convergence.

Motivated by the recent progress in compressing sensing, windowed l_1 relaxation is applied to the cost function of standard NLMS algorithm (Eq. 2) increasing the total number of near-zero coefficients, therefore reducing the computational load. Only coefficient with magnitudes in the range $[a - b]$ are attracted to the zero (Eq. 2), reducing steady-state misalignment. Upper and lower threshold of the attraction range are set after numerous runs to 0.0001 and $3.3354 \cdot 10^{-6}$ respectively, compromising between sparsity and accuracy of the fitting. Filter coefficients are update according to Eq. 2.

$$J = \sum_{k=1}^K e_k^2 + \sum_{k=1}^K \gamma_k \cdot \sum_{i=1}^M \sum_{j=1}^{N_p} |w(i, j)|$$

$$w_{n+1}(i, j) = w_n(i, j) + \sum_{k=1}^K \frac{\mu}{\sum_{i=1}^M x_k^2(i, j)} \cdot \left\{ x_k(i, j) \cdot e_k - \gamma_k \cdot \text{sgn}_w[w_n(i, j)] \right\} \quad (2)$$

$$\text{sgn}_w[w_n(i, j)] = \begin{cases} \text{sgn}[w_n(i, j)], & a < |w_n(i, j)| \leq b \\ 0, & \text{elsewhere} \end{cases}$$

with the parameter γ_k balancing between sparsity and estimation error.

WZA-NLMS is further extended to the algorithm that applies zero attraction only to the microphones that fail power criteria suggested by Eq. 3. In each iteration during the identification, microphone with the strongest

power of the impulse response is found ($\max\{P_{j_n}\}$) and all the microphones with the power bigger or equal to the half that value are considered significant and are excluded from the zero attraction. Advantage of such a approach is elimination of certain microphones from the array, reducing hardware costs.

$$P_{j_n} = \sum_{i=1}^M w_n^2(i, j) \quad j = 1 \cdots N_p \quad (3)$$

$$P_{criterion} = \begin{cases} \text{coeff. update eq. 2,} & \frac{P_{j_n}}{\max\{P_{j_n}\}} \cdot 100 \leq 50\% \\ \text{coeff. update eq. 1,} & \text{otherwise} \end{cases}$$

Finally, a modified RZA-NLMS algorithm is proposed, employing log-sum penalty to the cost function, leading to the re-weighted zero attractor in update equation (Eq. 4). Difference from basic RZA-NLMS introduced in (6) is that criteria for the selective shrinking is moved from the single filter tap to the complete impulse response of the microphone, where strong zero attraction takes place if the total some of absolute values of microphone taps is much less than ε . Otherwise, very little shrinking is applied to microphones classified as significant for the identification. ε is chosen as $(b - a) \cdot M$, with the same values for a , b and M as above for consistency and comparison. Compared to any ZA-NLMS algorithm that uses constant zero-attractor, proposed RZA-NLMS shows robustness for near-sparse systems.

$$J = \sum_{k=1}^K e_k^2 + \sum_{k=1}^K \gamma_k \cdot \sum_{j=1}^{N_p} \log \left(1 + \frac{\sum_{i=1}^M |w(i, j)|}{\varepsilon} \right) \quad (4)$$

$$w_{n+1}(i, j) = w_n(i, j) + \sum_{k=1}^K \frac{\mu}{\sum_{i=1}^M x_k^2(i, j)} \cdot \left\{ x_k(i, j) \cdot e_k - \frac{\gamma_k}{1 + \frac{\sum_{i=1}^M |w_n(i, j)|}{\varepsilon}} \cdot \text{sgn}[w_n(i, j)] \right\}$$

To further improve fitting performance, RZA-NLMS is extended to the PRZA-NLMS balancing between algorithm computational load and identification performance. Additionally, RW-RZA-NLMS is investigated, that uses roulette wheel rule for random selection of significant/insignificant microphones. Microphones with stronger impulse responses are with the higher probability chosen as significant. Advantage is in possibility to eliminate redundant microphones from the identification procedure.

3. MEASUREMENT SET UP

For testing the virtual sensing techniques and active noise control, a driver's cabin that simplifies the geometry but has the same overall dimensions $1.4 \text{ m} \times 1.4 \text{ m} \times 1.5 \text{ m}$ is constructed (Figure 2a). Aluminum alloys, wooden panels and plexiglas are used as building material. The cabin was gradually treated with porous absorbing material until the low frequency response at certain locations corresponded to the measurements performed in the cabin of a real agricultural machine. Commercial software for acoustical and vibration modeling (Actran) is used to perform modal extraction analysis of the constructed mock up cabin. Since the dominant noisy components related to the engine, cutter-head and blower of the machine are found to be in the frequency range from 80 Hz to 250 Hz, only the modes in this range are considered. By looking at the spatial pressure norm at modal frequencies around the target tonal components, the optimal number and the positions of the controlling sources are found. Three loudspeakers are placed on the roof of the cabin: at the left and right back corners, and at the front towards the center, as shown on Figure 2b (blue circles). For the noise control and virtual sensing measurements, a Head And Torso Simulator (HATS) Type 4128C from Brüel & Kjær with built-in left/right in-ear microphones is seated at the expected driver's place, mimicking realistic driving conditions. Forty six high quality microphones of the same type (ICP microphone, 4 mA, CAE Software and Systems GmbH) are randomly positioned on the cabin roof (Figure 2b) with somewhat higher density in the vicinity of the driver's head.

A high performance PXI based data acquisition system (DAQ) from National Instruments is used for synchronous sampling of the forty eight input channels. Channels are labeled from "0-47" with the channels "14" and "47" connected to the right/left ear of the HATS respectively. Microphones in the array on the roof are labeled as shown in Figure 2b. During the measurements, each secondary source (left, right and front) is excited, one at a time, with band-limited white noise (80 Hz-400 Hz), and signals from all microphones are

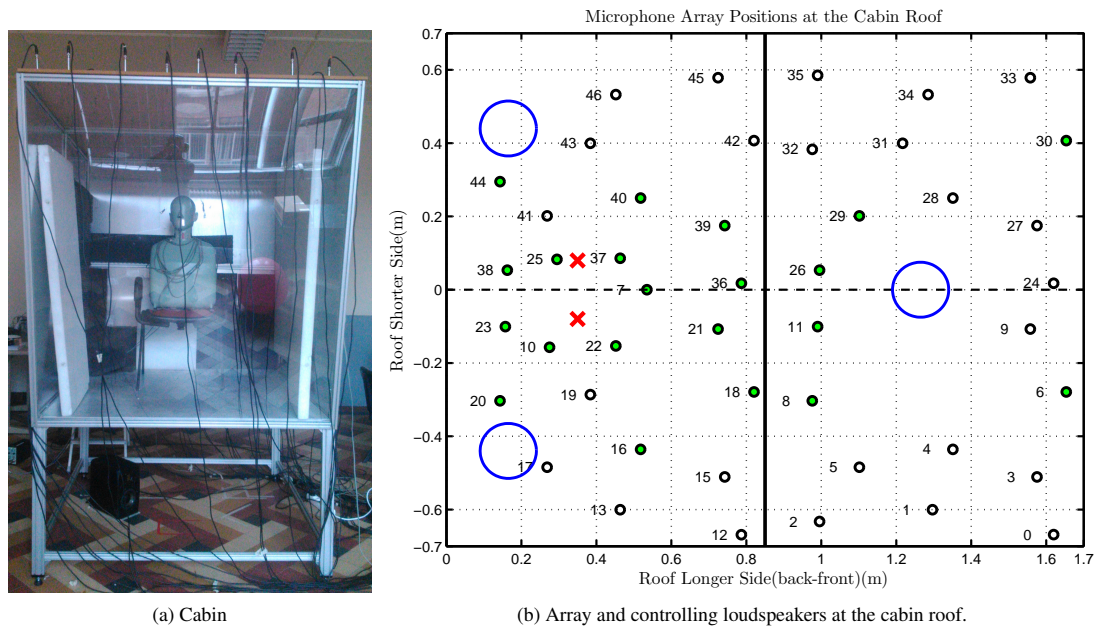


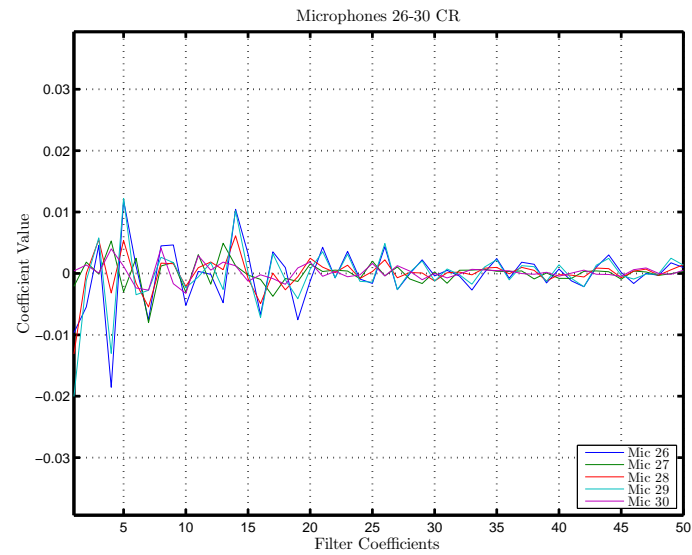
Figure 2 – Measurement setup.

simultaneously recorded. The same measurement procedure was repeated for the horizontal head rotations of 20, 45 and 60 degrees, considered to be angles of natural movements during the driving activity. Measurement based simulations are implemented in Matlab for the comparison of identification methods. They are compared

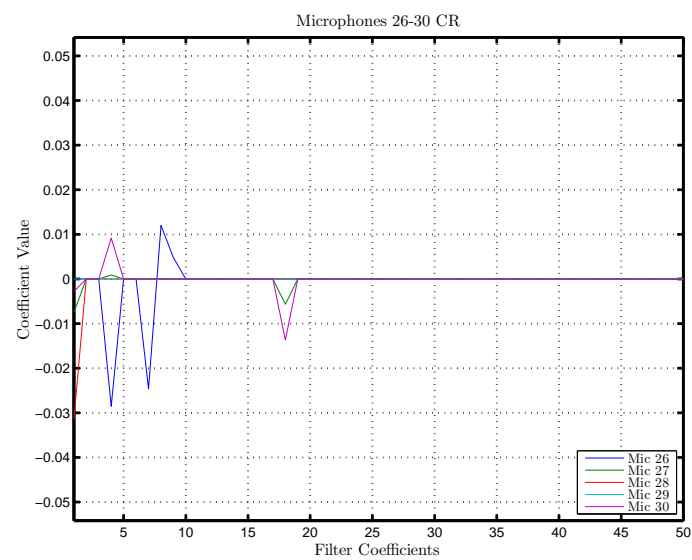
4. RESULTS

Favoring the sparsity is a compromise between fitting performance and system complexity. Applying a zero attractor to the filter coefficients increases the number of zero taps in the filters, but also leads to biased estimates. This is especially true for near sparse systems where not many coefficients are zero and is confirmed by the results reported here. Table 1 compares proposed identification methods in terms of fitting quality to the desired signal at the virtual location for the basic driver head position (head at 0 degrees in the horizontal plane), but also in terms of hardware cost required. As mentioned earlier, zero-shift cross-correlation between estimated and recorded signals is used as an indicator of estimation quality. Hardware costs are assessed with number of microphones chosen to be eliminated from the array ("Eliminated" column in Table 1). According to Table 1, the best fit is, as expected, achieved with reference the NLMS algorithm when all available forty six microphones were used. Even though WZA-NLMS algorithm promotes sparsity, it is not adequate for the target objectives of this paper, since it applies the same zero attractor to all microphone coefficients belonging to the pre-defined amplitude range, and is therefore unable to completely eliminate microphone from the identification procedure. Moreover, attracting non-zero coefficients to zero leads to a bigger mean square error, as it is confirmed by the fitting results of WZA-NLMS algorithm. Improved identification is achieved with the PWZA-NLMS algorithm, since some of the microphones were not influenced by zero attraction as their impulse responses passed defined power criteria. Converged impulse responses of these microphones remained complete. Moreover, 9/11 microphones turned to be insignificant for right/left ear signal fitting. However, with proposed parameters ($[a - b]$, power threshold), none of these microphones belonged to the right-left set intersection as they turned to be important for only one virtual location (ear). This disabled the possibility of completely removing these microphones from the roof. Figure 3 shows the influence zero-attractor had on converged filter coefficients for the same group of microphones (labeled as "26-30") when NLMS, WZA-NLMS and PWZA-NLMS are used. While NLMS keeps impulse responses complete, WZA-NLMS does it for all coefficient in the $[a - b]$ attracting range and PWZA-NLMS makes the selection between microphones rather than just filter taps.

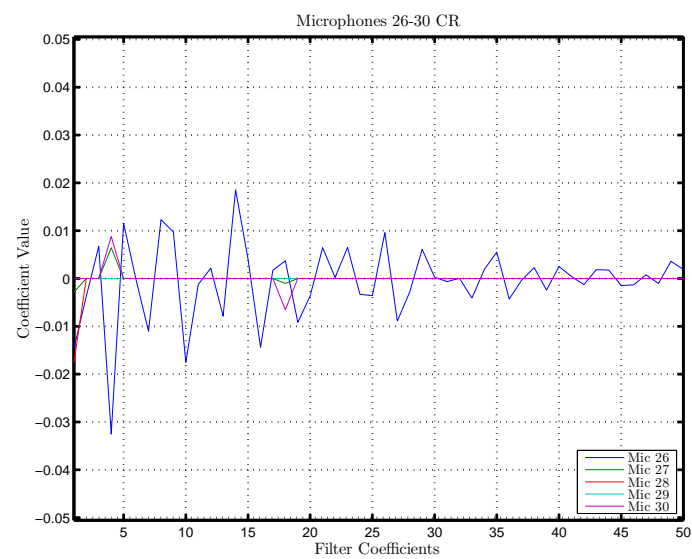
The proposed modification to the RZA-NLMS algorithm shows improved fitting for the right/left loudspeaker, as compared to the so-far the best PWZA-NLMS algorithm, while somewhat lower performance is achieved for the front loudspeaker. This difference might come from the fact that modified RZA-NLMS still applies zero attraction to all the coefficients, while attraction is significant for the microphones with less



(a) NLMS



(b) WZA-NLMS



(c) PWZA-NLMS

Figure 3 – Converged coefficients for microphones "26-30" for different identification methods. The virtual microphone is placed at the right ear.

Table 1 – Zero-shift cross-correlation based fitting achieved with different derivatives of LMS algorithm with HATS at "0" degrees.

Ear	NLMS		WZA-NLMS		PWZA-NLMS		RZA-NLMS		PRZA-NLMS		RW-RZA-NLMS	
	R	L	R	L	R	L	R	L	R	L	R	L
Front LP	97.44	97.77	85.11	84.88	94.61	94.87	93.64	93.90	93.87	94.58	94.41	94.97
Right LP	99.09	98.98	93.42	92.84	97.30	96.26	97.46	96.28	97.50	96.53	98.10	96.51
Left LP	98.86	99.10	95.23	93.38	96.31	97.35	96.71	96.34	96.97	97.48	96.95	97.75
Eliminated/Ear	0	0	0	0	9	11	24	26	24	27	22	25
Eliminated	0		0		0		13		15		12	

energetic impulse responses. These microphones are found to be very important for the front loudspeaker fit, which explains decrease in the performance. Selective shrinking of microphones with large and small sum of the coefficient's magnitudes outperformed the constant zero attraction of WZA-NLMS algorithm. Moreover, the number of insignificant microphones for the right/left ear is increased to 24/26, out of which 13 were common for both locations. As those microphones converged to the zero-valued impulse response for both left and right ear, they could be in the future removed from the roof.

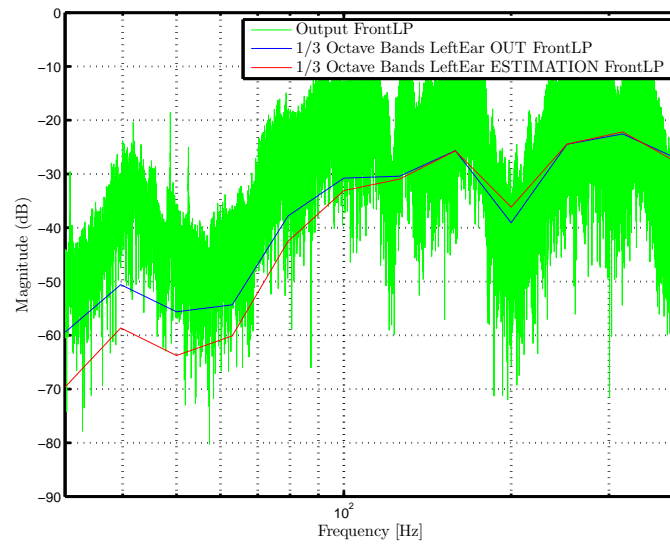
Since PRZA-NLMS applies RZA-NLMS algorithm only to the impulse responses of the microphones that are not passing the power criteria mentioned in Eq. 3, while remaining microphones are not affected with zero attraction, an increased performance is expected. Compared to the RZA-NLMS, both the fitting to for all three positions of the loudspeaker and the number of eliminated microphones are improved (Table 1). Second extension of basic RZA-NLMS algorithm also uses additional criteria to make an online selection of optimal microphones, but does it randomly with roulette-wheel rule. However, microphones with the stronger impulse responses are with the higher probability of being chosen. Related to this, fitting is improved for the front loudspeaker since in some iterations weaker microphones are selected, but on the other hand number of eliminated microphones decreased exactly for the same reason. Moreover, with utilizing random selection, RW-RZA-NLMS should be able to eliminate redundant microphones converging to the similar responses. This property was hard to be observed with performed analysis, but should be checked in the future.

Finally, NLMS algorithm is run only with the optimal microphones derived with RZA-NLMS method, since all the other methods give more-less similar solutions. All microphones that had the power less than $\frac{1}{10} - th$ the power of the strongest microphone were eliminated. Of course, those with a zero power were eliminated with RZA-NLMS during the identification procedure. Twenty one microphones remained and they are marked as a "green circles" on Figure 2b. Figure 4 shows fitting results in 1/3-octave bands (31.5 Hz-400 Hz) for the left ear and all three loudspeaker positions. For the ANC frequency band of interest (100 Hz-250 Hz), achieved results seem acceptable.

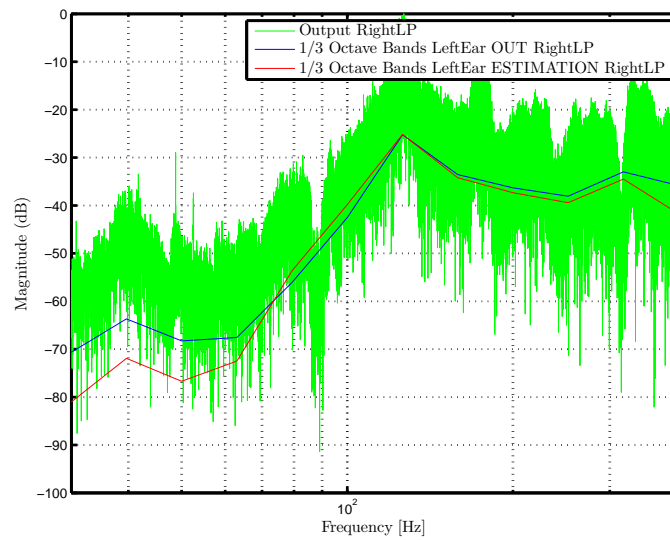
For the sweet spot analysis reference NLMS algorithm is compared to the RW-RZA-NLMS and results are presented in the Table 2. Intention was to asses sensitivity of each algorithm to the head rotations, but also understand how reduction of microphones copes with maintaining the sweet spot in comparison to the reference case. For both algorithms, front loudspeaker loses the sweet spot the fastest, already for more than 20 degrees of rotation. On the other hand, sweet spot for the left and right loudspeaker is satisfying even for the rotations of 45 degrees and probably is acceptable for 60 degrees as well, although this should be tested in the application itself. Interesting observation is improvement in sensitivity to the head movements when number of microphones is reduced. This confirms initial assumptions on possible over-fitting issues of NLMS approach.

Table 2 – Sweet spot evaluation.

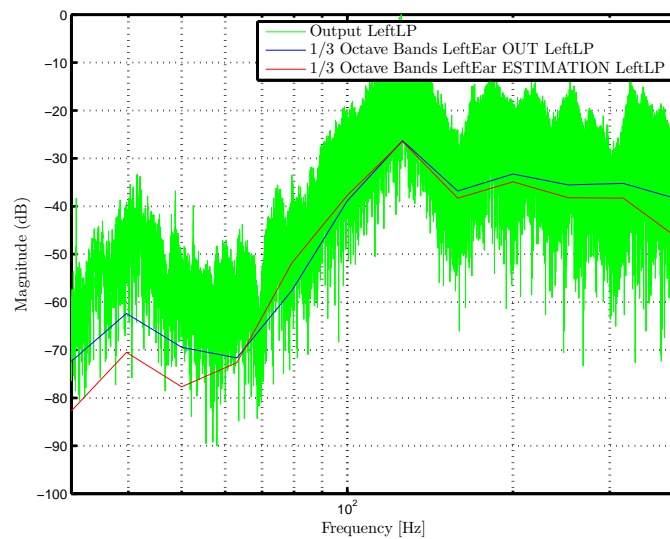
Algorithm		HATS-0		HATS-20		HATS-45		HATS-60	
		R	L	R	L	R	L	R	L
NLMS	Front LP	97.44	97.77	95.04	95.76	91.15	94.51	85.64	89.31
	Right LP	99.09	98.98	99.07	98.81	98.34	97.03	97.87	95.83
	Left LP	98.86	99.10	98.64	98.80	97.19	97.35	95.76	95.58
RW-RZA-NLMS	Front LP	94.41	94.97	92.00	92.84	88.22	92.32	82.28	86.63
	Right LP	98.10	96.51	98.03	95.89	97.43	93.44	97.07	92.23
	Left LP	96.95	97.75	97.20	97.49	96.64	96.23	95.43	94.56



(a) Front LP-Left Ear



(b) Right LP-Left Ear



(c) Left LP-Left Ear

Figure 4 – Fitting results in 1/3-octave bands (31.5 Hz-400 Hz) for three controlling source positions with the NLMS estimation and optimal number of microphones. Total spectrum at left ear (green), 1/3-octave band spectrum at left ear (blue), 1/3 octave band spectrum left ear estimated (red).

5. CONCLUSIONS

Virtually sensing the sound level at the ear of an operator of a machine based on a remote array of microphones could be useful for checking the error in applications such as active noise control. In this paper the methodologies for obtaining the filters needed to extract the virtual microphone signals that were previously proposed have been applied to the low frequency components in a driver's cabin. In this resonant cabin, placing microphones as close as possible to the head is not necessarily the best control strategy. Hence, starting from a dense, random array of microphone positions located in the roof, adapted algorithms were proposed that at the one hand eliminate as many of the filter taps as possible to reduce the computational load, and at the other hand eliminate as many microphones as possible. It was shown that these algorithms indeed reduce the set of microphones in a way that could not be foreseen on the basis of correlation coefficients between array microphones and the microphones placed at the virtual locations during training. The stability of the resulting solution was maintained and thus the solution generalizes well, as could be seen from an analysis of the sweet spot when a HATS is turned slightly away from the original position.

REFERENCES

1. Faber BM, Sommerfeldt SD. Global control in a mock tractor cabin using energy density. In: Proceedings of the 2004 International Symposium on Active Control of Sound and Vibration (ACTIVE '04). Williamsburg, Virginia, USA; 2002.
2. Cazzolato B. An adaptive LMS virtual microphone. In: Proceedings of the 2002 International Symposium on Active Control of Sound and Vibration (ACTIVE '02). Southampton, United Kingdom; 2002.
3. Moreau D, Cazzolato B, Zander A, Petersen C. A review of virtual sensing algorithms for active noise control. *Algorithms*. 2008;1(2):69–99. Available from: <http://www.mdpi.com/1999-4893/1/2/69>.
4. Jin J, Qu Q, Gu Y. A robust zero-point attraction LMS algorithm on near sparse system identification. arXiv preprint arXiv:13032255. 2013.
5. Petersen CD. Optimal spatially fixed and moving virtual sensing algorithms for local active noise control. PhD Thesis. School of Mechanical Engineering, The University of Adelaide; 2007.
6. Chen Y, Gu Y, Hero A. Sparse LMS for system identification. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Taipei, Taiwan; 2009. p. 3125–3128.